Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 4366–4374
Bioorganic & Medicinal Chemistry

Theoretical quantitative structure–activity relationships of flavone ligands interacting with cytochrome P450 1A1 and 1A2 isozymes

F. Iori,[a] R. da Fonseca,[b] M. João Ramos[b] and M.C. Menziani[a,*]

[a]*Dipartamento di Chimica, Università degli Studi di Modena e Reggio Emilia, Via Campi 183, 41100 Modena, Italy*
[b]*Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169–007 Porto, Portugal*

Received 19 November 2004; accepted 22 April 2005
Available online 23 May 2005

Abstract—Theoretical descriptors obtained from quantum mechanical calculations on isolated ligands in different media and molecular dynamics simulations of ligand–enzyme complexes have been used to obtain a quantitative rationalization of the inhibition of CYP1A2 and CYP1A2 by three series of flavonoids. Predictive models obtained through one-descriptor QSAR studies and mechanistic explanations have been obtained for recognition and selectivity.

1. Introduction

Ubiquitously present in fruits and vegetables, flavonoids are important constituents of the human diet and are actively investigated as possible chemopreventive agents against environmental mutagens and carcinogens.[1]

One important mechanism by which flavonoids exert this effect is the inhibition of the 1A1 and 1A2 cytochrome P450 (CYP) isozymes.[2,3] In spite of the high amino acid sequence identity (72%), the two isozymes display different substrate specificity and inhibitor susceptibilities. CYP1A1 mainly metabolizes polycyclic aromatic hydrocarbons to their toxic derivatives[4] and is constitutively expressed in many tissues such as lung, gastro-intestinal tract, placenta, brain, and in vascular endothelial and smooth muscle cells; CYP1A2 preferentially oxidizes heterocyclic and aromatic amines[4] and is expressed constitutively in hepatic tissues.

The understanding of the exact mechanism of flavonoid actions in the human body is limited since, depending upon their structures, concentrations, and experimental conditions, flavonoids can either inhibit or activate human CYPs. Moreover, they can be metabolized by several CYPs, giving rise to metabolites with associated biological activities distinctly different from those of the parent compounds.[1] In addition, the data available are mainly obtained by means of in vitro experiments. Far fewer animal studies have been carried out and the results obtained may not be predictive of effects in humans because of differences in metabolism of flavonoid compounds between species.[5] To complicate the matter, the molecular series analyzed in each laboratory generally contain small numbers of compounds and the different experimental conditions used in different laboratories yield inhomogeneous data, often contradictory.

Therefore, it would be of great interest to elaborate a tool able to cast some light onto the structural basis responsible for the differences in the mode of action of this class of compounds.

Structure–function relationship studies on the inhibitory effects of flavonoids toward CYP1A1 and CYP1A2 carried out over the last few years have been recently reviewed,[1] and some quantitative correlations have been attempted.[6–9] In these studies, predictive models and/or mechanistic explanations have been obtained by means of multi-parameter approaches mainly based on Molecular Orbitals (MO) derived descriptors.

Additional information on the molecular determinants for flavonoid action and selectivity can be obtained by docking the ligands into the active site of the three-dimensional structures of the isozymes. In the absence of experimental structural data, several homology models based on the crystal structures of bacterial enzymes, and more recently, mammalian CYPs, have been derived for the 1A1 or the 1A2 isozymes.[10–13]

* Corresponding author. Tel.: +39 059 2055091; fax: +39 059 373543; e-mail: menziani@unimo.it

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.04.066

In this study, quantitative structure–activity relationship (QSAR) models have been derived for three series of flavonoids[6,14,15] by making use of descriptors obtained from quantum mechanics calculations on the isolated inhibitors and from molecular mechanics calculations on the ligand–enzyme complexes. The models can be used as tools to extract information on the binding and selectivity determinants of structurally related compounds from heterogeneous experimental data measured in different laboratories.

2. Results and discussion

A general limitation in the computational rationalization of CYP inhibition selectivity by flavonoids is the availability of experimental data on selective ligands in the current literature, such that a numerically consistent set of compounds can be investigated. In fact, the inhibitory capacity of flavonoids with respect to CYP activities has been extensively studied by several laboratories,[1] but the different experimental protocols used do not allow a direct comparison of the data obtained. The data we use in this study have been measured by Lee et al.[6] (first series of compounds, listed in Figure 1) and by Zhai et al.[14,15] in two concomitant studies (second and third series of compounds, listed in Figure 2). The first study concerns the inhibitory effects of flavonoids on the caffeine N3-demethylation (C3ND) activity of CYP1A2 from human liver microsomes. The data are given as IC_{50} values. The second study focuses on the effects of flavone and five hydroxylated derivatives on the methoxyresorufin *O*-demethylase (MROD) activity catalyzed by cDNA-expressed human cytochromes P450 CYP1A1 and CYP1A2 ($\log 1/IC_{50}$), and the third examines a series of 10 structurally related flavonoids, partially overlapping with the second study, whose inhibitory effect, assessed on human liver microsomes at two concentrations (0.15 and 0.5 μM) and expressed as % of control (MROD), was considered to be due to CYP1A2.

The lack of large data sets of homogeneous experimental data requires the derivation of quantitative mono parametric models based on highly informative descrip-

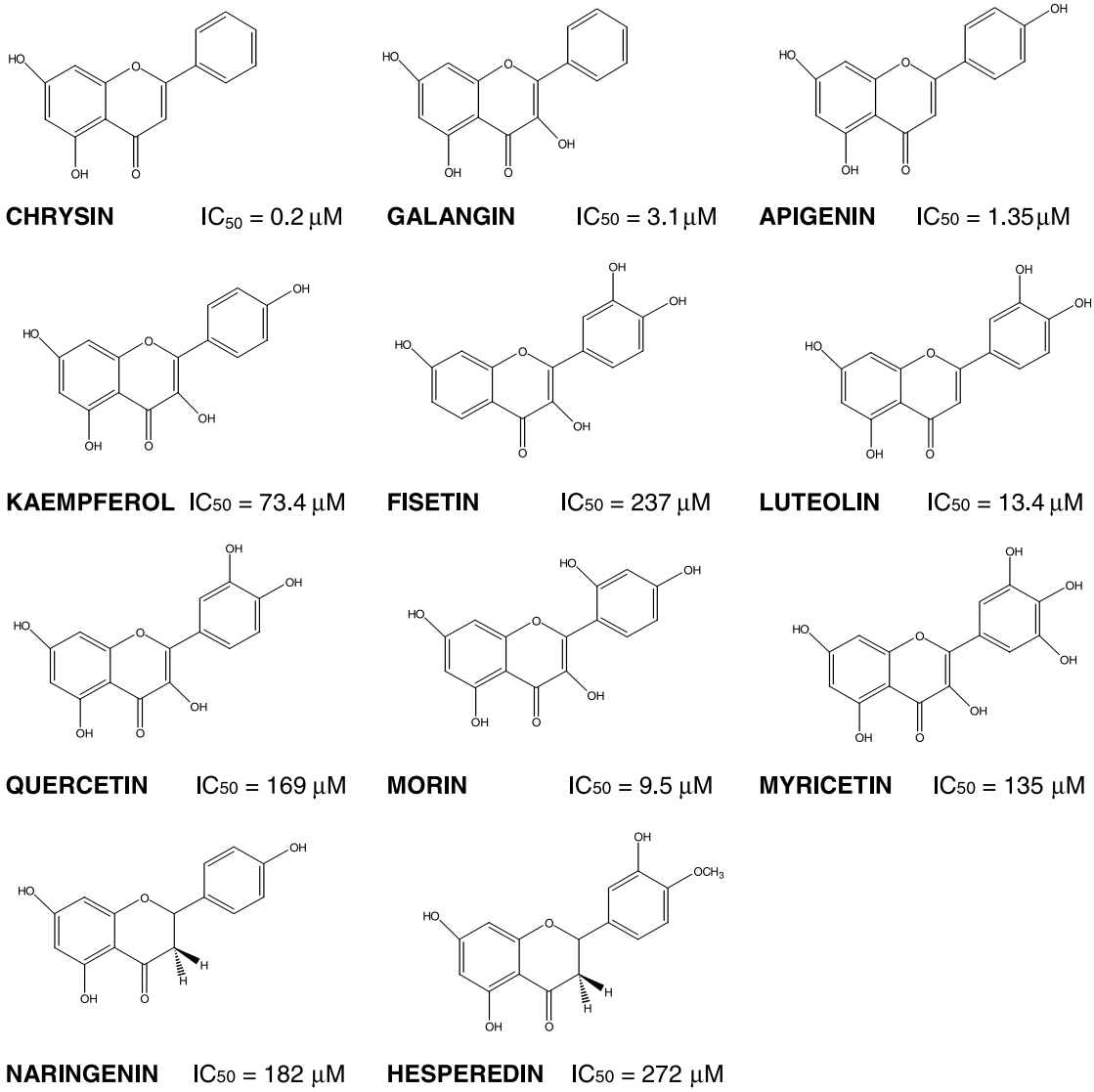
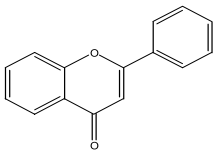
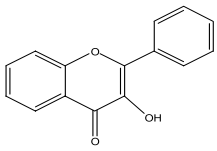


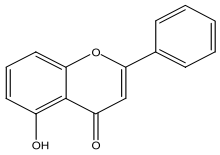
Figure 1. Molecular structure and CYP1A2 activity of flavonoids constituting the first series of compounds studied.



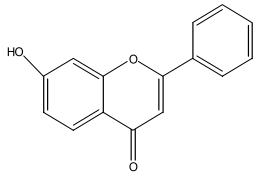
FLAVONE
IC_{50}(CYP1A1) = 0.14μM; IC_{50}(CYP1A2) = 0.066μM
% ATT(0.15 μM) = 46.87; %ATT(0.50 μM) = 20.3



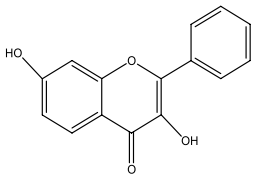
3-OH-FAVONE
IC_{50}(CYP1A1) = 0.035μM;IC_{50}(CYP1A2) = 0.026μM
%ATT(0.15 μM) = 45.3; %ATT(0.5 μM) = 17.18



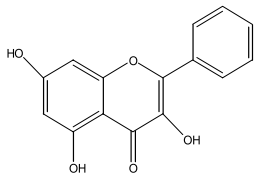
5-OH-FLAVONE
IC_{50}(CYP1A1) = 0.066μM; IC_{50}(CYP1A2) = 0.026μM
%ATT(0.15 μM) = 56.25; %ATT(0.50 μM) = 28.12



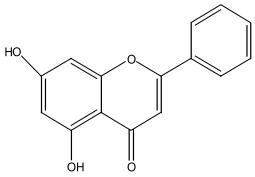
7-OH-FLAVONE
IC_{50}(CYP1A1) = 0.035μM;IC_{50}(CYP1A2) = 0.222μM
%ATT(0.15 μM) = 78.12; %ATT(0.50 μM) = 62.5



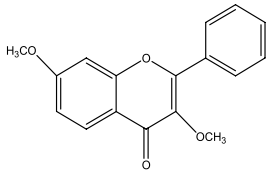
3,7-diOH-FLAVONE
IC_{50}(CYP1A1) = 0.04μM; IC_{50}(CYP1A2) = 0.111μM
%ATT(0.15 μM) = 67.18; %ATT(0.50 μM) = 39.06



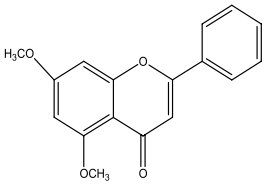
3,5,7-triOH-FLAVONE (GALANGIN)
IC_{50}(CYP1A1) = 0.095μM; IC_{50}(CYP1A2) = 0.018μM
%ATT(0.15 μM) = 54.68; %ATT(0.50 μM) = 29.68



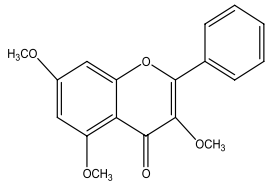
5,7 -diOH-FLAVONE (CHRYSIN)
%ATT(0.15 μM) = 18.75; %ATT(0.50 μM) = 9.38



3,7-diOCH_3-FLAVONE
%ATT(0.15 μM) = 50.0; %ATT(0.50 μM) = 31.25



5,7-diOCH_3-FLAVONE
%ATT(0.15 μM) = 84.38; %ATT(0.50 μM) = 68.75



3,5,7-triCH_3-FLAVONE
%ATT(0.15 μM) = 67.19; %ATT(0.50 μM) = 40.62

Figure 2. Molecular structure and activities of flavonoids constituting the second and third series of compounds studied.

tors. The ones used in this paper are quantum mechanical descriptors computed on the isolated molecule in water solution and low dielectric solvent and molecular mechanics descriptors computed on three dimensional models of the ligand–receptor complexes.

Among the large number of descriptors computed, we have selected the ones that furnish the most significant correlations, both from a statistical and an interpretative point of view, and are common to the three series of compounds studied. Table 1 lists the selected descriptors for the first series of ligands, together with their inhibition constants normalized with respect to the galangin derivative ($pIC_{50} = -\ln(IC_{50X}/IC_{50Galangin})$). They are: (a) the CHARMM[16] total interaction energy between the ligand and the enzyme (IE_{tot}). In order to

Table 1. Molecular descriptors and activity parameters for the first series of compounds studied

Ligands	pIC_{50} CYP1A2	IE/Surf (kcal/mol Å^2)	IE_{HEME} (kcal/mol)	IE_{HB} (kcal/mol)	$HB_{acceptor}$	$\Delta(\delta G_s)_{ZINDO}$ (kcal/mol)
Chrysin	2.728	−0.463	−10.72	−1.37	20.59	−0.160
Apigenin	0.818	−0.454	−9.15	−2.73	24.73	0.007
Galangin	0.000	−0.443	−8.31	−3.94	25.52	−0.046
Morin	−1.129	−0.349	−6.61	−3.56	27.6	0.219
Luteolin	−1.477	−0.363	−6.73	−6.06	27.26	0.135
Kaempferol	−3.178	−0.273	−4.96	−10.08	31.45	0.237
Quercetin	−4.011	−0.256	−2.03	−12.83	30.59	0.287
Naringenin	−4.086	−0.300	−1.89	−10.51	28.03	0.229
Myricetin	−4.102	−0.315	−2.08	−13.26	31.34	0.303
Fisetin	−4.350	−0.230	−1.26	−11.57	32.28	0.115
Hesperetin	−4.487	−0.247	−1.00	−12.37	27.75	0.306

render this index comparable among different series of compounds it has been computed per unit surface of the ligand (IE/Surf). In fact, the main contribution to IE_{tot} is given by the 6–12 nonbonded dispersion and repulsion terms, which is proportional to the size of the molecule; (b) the CHARMM total interaction energy between the ligand and the heme, which is mainly of electrostatic nature (IE_{HEME}); (c) the CHARMM hydrogen bonding contribution to the total interaction energy between the ligand and the enzyme (IE_{HB}); (d) the hydrogen bond acceptor complementarity between the protein and the ligand ($HB_{acceptor}$), computed with the METASITE program;[17] and (f) the ZINDO[18] free energy of desolvation, i.e. the energy cost of transferring ligands from water into the binding site ($\Delta(\delta G_s)$).

Because of their high intercorrelations (Table 2) these descriptors can only be used in a monovariate analysis for predictive purposes. The attempt to capture and interpret the decisive determinant for inhibition is hindered, but different events that concur to the formation of the ligand–receptor complex can be highlighted.[19]

As expected, the inhibitor potency of the ligands is directly related to the stability of the complex it is able to form (Fig. 3A). This might be considered an obvious result, but what is not straightforward is that the most active compounds shows the most stable interactions with the heme (Fig. 3B) and the weakest hydrogen bond interaction energy with the protein (Fig. 3C). The latter trend is also confirmed by the slope of the correlation obtained with the $HB_{acceptor}$ descriptor (Fig. 3D), which indicates that a weaker complementarity between the ligand hydrogen bond donor ability and the receptor hydrogen bond acceptor propensity is required for high inhibitory potency. This correlation is notably improved by omitting the structurally more dissimilar flavanone compounds (Naringenin and Hesperedin) from the regression. In this case, the new correlation obtained yields a correlation coefficient $R^2 = 0.97$.

The analysis of the trajectories derived from the molecular dynamics study of the complexes helps to explain this requirement. In fact, the ligands carrying several hydroxyl groups get trapped into a polar pocket in proximity to the heme site, leaving unperturbed the catalytic site and the associated network of water molecules. As an example, a zoom into the binding site of the minimized average structure of the Quercetin–*h*CYP1A2 complex is shown in Figure 4. The hydrogen-bond network formed with residues F186, T458, T84, and D273 cannot be broken in the time-scale of the dynamics. It is worth noting that of particular importance for the realization of such a strong network is the simultaneous presence of the hydroxyl groups in position 3 and 3′. A similar behavior is, in fact, observed during the dynamics of Fisetin and Myricetin, whereas the steric interaction between the hydroxyl group in position 2′ of Morin and F186 prevents the attraction and permanence of the compound in the polar pocket. Ligands presenting structures less crowded by hydroxyl substituents accommodate deeper into the binding site, with the pendant phenyl ring almost perpendicular to the heme, as shown in Figure 4 for the Kaempferol–*h*CYP1A2 complex. Indeed it has been shown[1,3] that Galangin is metabolized by CYP1A1/1A2, being sequentially transformed to Kaempferol and then Quercetin, which is responsible for Galangin mutagenicity in the *Salmonella typhimurium* test. Therefore, the two different binding modalities hypothesized on the basis of the dynamics behavior of these compounds in interaction with CYP1A2 are consistent with the fact that Galangin and Kaempferol are contemporaneous inhibitors, suppressing the genotoxicity of chemicals, and substrates acting as harmful prooxidant compounds.

Table 2. Correlation matrix (correlation coefficients (R) are listed) of the data listed in Table 1

	pIC_{50} CYP1A2	IE/Surf	IE_{HEME}	IE_{HB}	$HB_{acceptor}$	$\Delta(\delta G_s)_{ZINDO}$
pIC_{50} CYP1A2	1.000					
IE/Surf	−0.941	1.000				
IE_{HEME}	−0.972	0.926	1.000			
IE_{HB}	0.955	−0.898	−0.950	1.000		
$HB_{acceptor}$	−0.985	0.929	0.939	−0.907	1.000	
$\Delta(\delta G_s)_{ZINDO}$	−0.880	0.820	0.829	−0.814	0.854	1.000

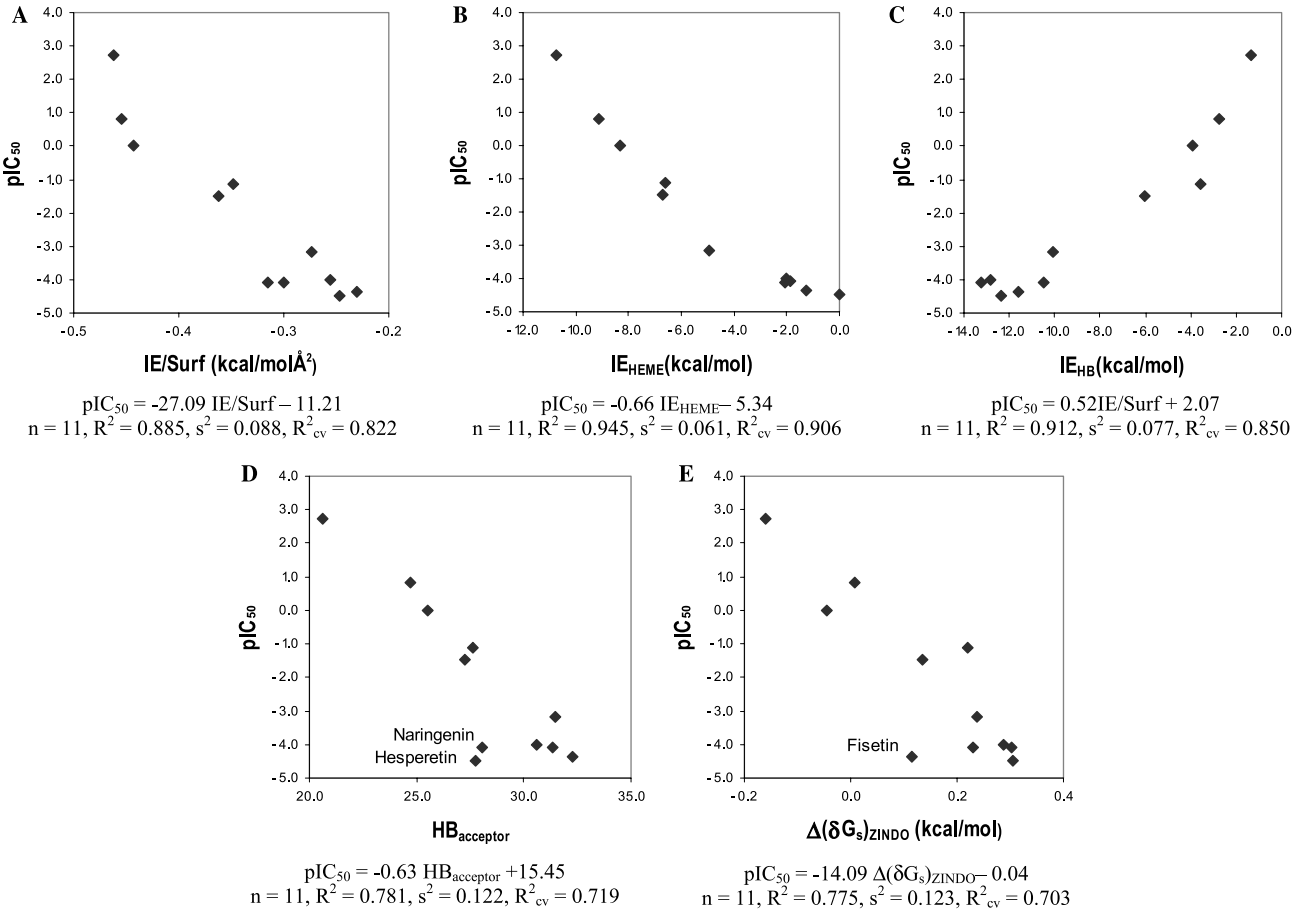


Figure 3. Selected correlations for the first series of ligands (CYP1A2 activity). Statistical parameters R = correlation coefficient, n = number of compounds, s = standard deviation are given. Assessment of the prediction power of the QSAR models is given by the cross-validated correlation coefficient R^2_{cv}.

Finally, the correlation obtained between the inhibitory activity and the free energy of desolvation (Fig. 3E) shows that the most active compounds are more stable in a low dielectric environment, so that a lower energy cost to desolvate them is necessary. Fisetin is overestimated by the equation, showing a $\Delta(\delta G_s)$ value very similar to Luteolin. Omission of this compound in the regression improves the coefficient correlation to $R^2 = 0.90$.

The correlations obtained for the second series are reported in Figure 5 and the corresponding data values are listed in Table 3. Notwithstanding the limited number of compounds available, satisfactory correlations are obtained both for the CYP1A2 and CYP1A1 inhibitory activity data. The IE/Surf descriptor explains 89% of the variance of the CYP1A2 inhibitory potency (Fig. 5A), whereas in the case of the CYP1A1 inhibitory potency the descriptor is able to cluster the data into two groups of comparable activity (Fig. 5C). Probably because of the lack of hydroxyl substituents on the pendant phenyl ring, in this series of ligands the inhibitory activity toward the two isozymes is directly proportional to the ability to form hydrogen bonds (Fig. 5B and D). Moreover, a very good correlation is observed between the CYP1A1 inhibitory activity and the desolvation free energy.

The third series of ligands studied partially overlapped with the second but is enriched by three methylated derivatives and the 5,7-dihydroxyflavone. The two series have been studied for CYP1A2 inhibitory activities in the same laboratory.[14,15] However, the data are not directly comparable and the best correlation coefficient among the measured activities ($R^2 = 0.85$) is obtained, for the compounds shared by the two series, between the inhibitory activities measured on cDNA-expressed CYP1A2 (pIC_{50}, Table 3) and human liver microsomal CYP1A2 at 0.15 μM of ligand concentration (pActivity, Table 4). However, a significant dispersion of the data is observed ($s^2 = 0.27$), especially for the most active ligands.

It has to be emphasized that in similar situations the correlations obtained can be judged only with respect to the intrinsic nature of the pharmacological data available. In this case, binding measures on an individual enzyme subtype (cloned enzyme) or on a generic population of enzymes where the one of interest is assumed to be predominant, but interactions between the subtypes cannot be ruled out.

Accordingly, poorer correlations are obtained for this series of compounds (see Fig. 6 and Table 4 for the

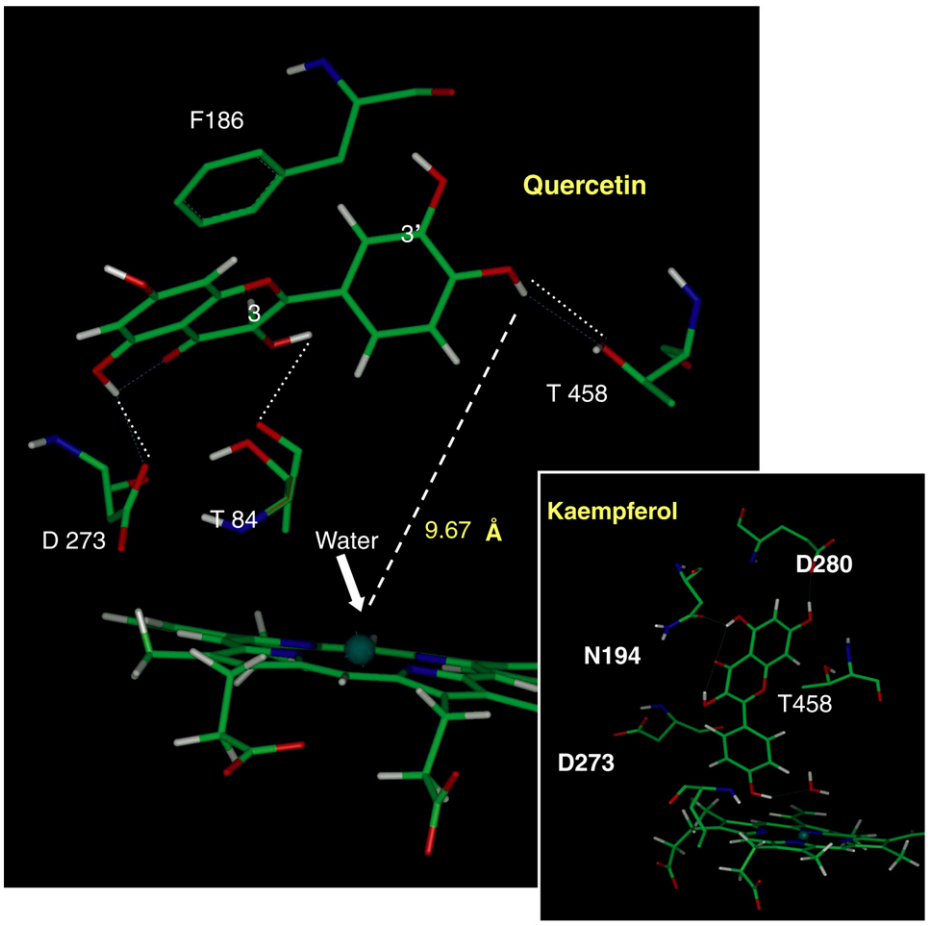


Figure 4. Zoom into the binding site of the of the Quercetin–*h*CYP1A2 and Kaempferol–*h*CYP1A2 complexes. The minimized average structures are shown.

corresponding descriptors data values). In the correlation obtained between pActivity (pActivity = $-\ln(\text{Activity}_X/\text{Activity}_{\text{Galangin}})$) and the IE/Surf descriptor, compound 3,5,7-OCH_3 is largely overestimated by the regression; in fact its omission yields a correlation coefficient $R^2 = 0.80$. It is plausible that to accommodate this bulky compound the binding site undertakes an energetically expensive conformational adjustment that is permitted in our dynamical study, but not accounted for by the descriptor used. In fact, in the IE/Surf descriptor the interaction energy IE_{tot} is simply computed by the equation $IE_{tot} = E_{[ligand-enzyme]} - E_{ligand} - E_{enzyme}$. Here $E_{[ligand-enzyme]}$ is the energy of the final minimized average structure of the complex, and E_{ligand} and E_{enzyme} are the energy of the ligand and of the enzyme in the conformation they assume in the complex. Therefore, the energy cost of the mutual conformational changes of the ligand and enzyme upon binding is not considered. The $HB_{acceptor}$ descriptor performs quite well, although a significant dispersion of the data is observed with respect to the corresponding correlation obtained for the second series of compounds. Moreover, flavone in this case is underestimated by the regression, whereas it was well predicted in the previous case (Fig. 5B).

A very interesting correlation is obtained for this series of compounds by means of the energy of the highest occupied molecular orbital (E_{HOMO}) of the ligands in water solution. The most active compounds show the most stable HOMO in water; therefore they are less reactive with respect to an electrophilic attack and this facilitates their dissolution in a lipophilic phase (enzyme).

Additional information on the molecular determinants for inhibitory activity and selectivity between the two isozymes can be obtained from the analysis of the minimized average structure of the complexes. The interaction energies between the ligands of the second series and individual residues of *h*CYP1A2 and *h*CYP1A1 are reported in Tables 5 and 6, respectively. Residues, which have been shown to be important for ligand binding by site directed mutagenesis studies[20–22] are reported in bold in the tables. In both isozymes the residues **F**186, **D**273/275, and **I**346/348 generally give strong interactions with all the derivatives and therefore seem to be responsible for ligand–enzyme recognition. The modulation of the ligand inhibitory activity seems to be mainly due to residues L71, S74, **V187**, N194, **N272**, D280, and T345 in CYP1A2 and residues **S84**, **G187**, S192, **D282**, and T347 in CYP1A1. Moreover, residues that strongly interact with ligands in CYP1A2 but not in CYP1A1 and vice versa might be considered responsible for the selectivity between the two isozymes. These are residues

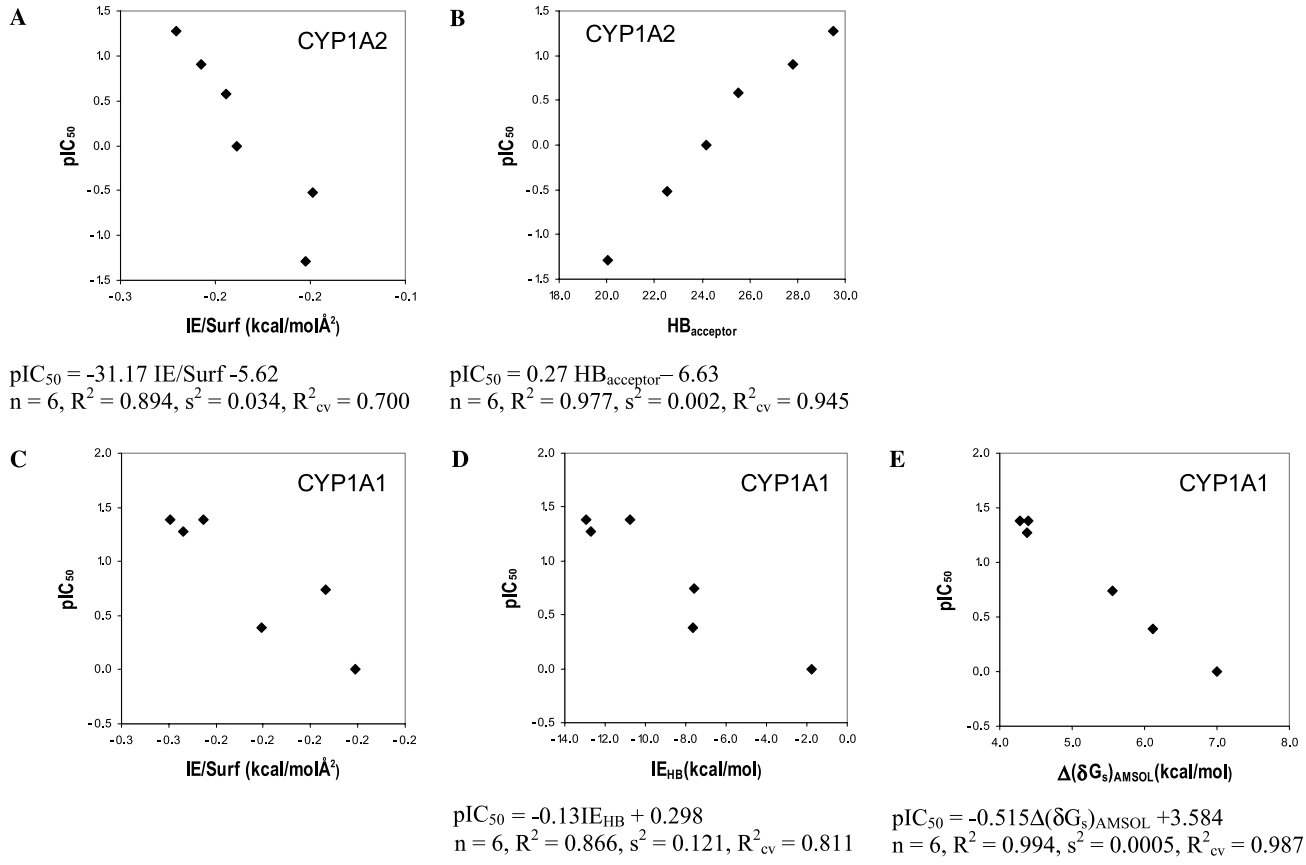


Figure 5. Selected correlations obtained for the second series of ligands (CYP1A2 and CYP1A1 activities). Statistical parameters R = correlation coefficient, n = number of compounds, s = standard deviation are given. Assessment of the prediction power of the QSAR models is given by the cross-validated correlation coefficient R^2_{cv}.

R68 and **T458** for CYP1A2 and residue **Y72** for CYP1A1.

3. Conclusions

Correlations for the inhibition of CYP1A1 and CYP1A2 by flavonoids have been derived in this study by means of theoretical descriptors obtained from quantum mechanical calculations on the isolated ligands in different media and molecular dynamics studies of the ligand–enzyme complexes. The lack of large enough experimental data sets dictates the need of highly informative parameters to be used for one-descriptor QSAR studies. These, together with the analysis of the molecular dynamics trajectories of the complexes, have been shown to be able to (a) identify different binding modalities for the ligands depending on the number and positions of the hydroxyl substituents they carry, (b) relate them with differences in the metabolism of anticarcinogens giving rise to metabolites with higher or lower biological activity than the parent compound, (c) provide a quantitative rationalization of the selectivity of a set of flavonoid ligands toward the CYP1A1 and CYP1A2 isozymes, and (d) suggest amino acids to be studied by site directed mutagenesis as putatively involved in ligand–enzyme recognition, modulation of the ligand inhibitory activity, and selectivity toward the isozymes.

4. Methods

Homology models of human cytochromes P450 1A1 and 1A2 have been constructed based on the structures

Table 3. Molecular descriptors and activity parameters for the second series of compounds studied

Ligands	CYP1A2			CYP1A1			
	pIC_{50}	IE/Surf (kcal/mol Å^2)	$HB_{acceptor}$	pIC_{50}	IE/Surf (kcal/mol Å^2)	IE_{HB} (kcal/mol)	$\Delta(\delta G_s)_{AMSOL}$ (kcal/mol)
Flavone	0	−0.189	24.17	0	−0.171	−1.74	7.00
3-OH	0.906	−0.207	27.83	1.384	−0.250	−12.92	4.39
5-OH	0.578	−0.194	25.5	0.742	−0.184	−7.61	5.55
7-OH	−1.29	−0.153	20.04	1.384	−0.235	−10.75	4.28
3,7-OH	−0.52	−0.149	22.53	1.273	−0.244	−12.73	4.37
3,5,7-OH	1.271	−0.221	29.52	0.386	−0.211	−7.66	6.12

Table 4. Molecular descriptors and activity parameters for the third series of compounds studied

Ligands	pActivity (0.15 mM)	IE/Surf (kcal/mol Å^2)	$HB_{acceptor}$	E_{HOMO} (eV)
Flavone	0	−0.189	0.110	−87.20
3-OH	0.066	−0.207	0.149	−87.57
5-OH	−0.177	−0.194	0.130	−86.71
7-OH	−0.518	−0.153	0.109	−85.51
3,7-OH	−0.351	−0.149	0.112	−86.08
5,7-OH	−0.177	−0.200	0.120	−86.62
3,5,7-OH	0.488	−0.226	0.162	−89.52
3,7-OCH_3	−0.405	−0.155	0.107	−84.91
5,7-OCH_3	−0.351	−0.176	0.115	−85.27
3,5,7-OCH_3	−0.575	−0.184	0.082	−83.94

of three templates: CYPterp (PDB entry: 1cpt), CYP-Bm3 (PDB entry: 1bu7), both from bacterial organisms, and the chimeric CYP2C5-2C3 (PDB entry: 1dt6) from rabbit, the first mammalian P450 to be crystallized. Details of the procedure used were reported previously.[10,12]

The sequences of human and rat microsomal cytochromes CYP1A1/2 were extracted from SWISS-PROT protein sequence database and evaluated with the PROF (http://cubic.bioc.columbia.edu/predictprotein) tool of secondary structure prediction present at PredictProtein@Embl-Heidelberg.de.[23] The residues corresponding to the membrane anchor N-terminal helix were ignored while building the model, as there is no template structure available for it. Each target sequence was aligned to those of the templates using ClustalW (http://www.ebi.ac.uk/clustalw/)[24] and the cytochromes models have been built by means of restraint-based comparative modeling, with the program Modeller.[25] Energy minimization and molecular dynamics were performed using the Charmm24[16] program with united atoms. Polar hydrogens were added to the models and the heme was fitted into each active site in analogy with its position in the templates.

The minimization procedure consists of 500 steps of steepest descendent and a number of steps of conjugate gradient sufficient to obtain a potential energy rms less than 1.0E−5 kcal/(mol Å). A dielectric constant $\varepsilon = 80$ has been used for the first minimization, then a second energy refinement was performed by using an explicit 8 Å layer of water molecules.

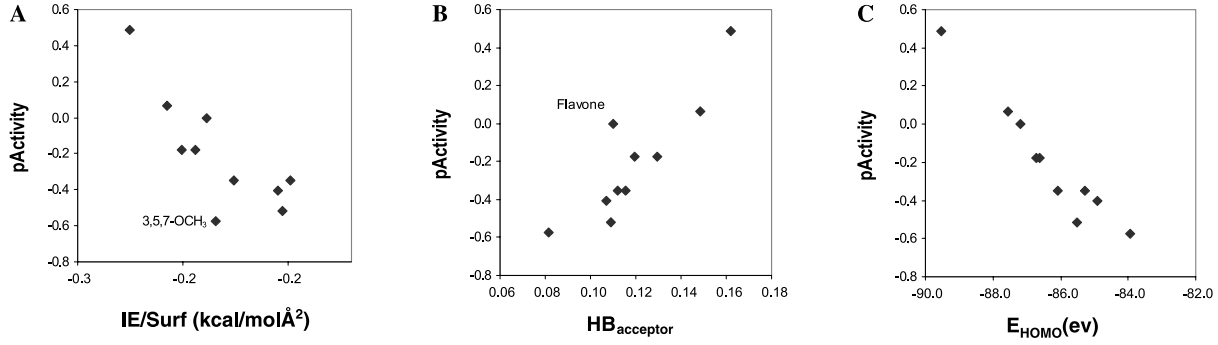
The minimized structures were dried and subjected to a 400 ps molecular dynamic simulations run with implicit solvent. The SHAKE option[26] was used allowing a 2 fs integration time step.

The technical quality of the models obtained has been evaluated with respect to the templates and the new X-ray resolved structure of human cytochromes 2C9.[27] To this purpose WhatIF,[28] Profile3D,[29] and Procheck[30] programs have been used. Briefly, the WhatIF score is a measure of the agreement between the distribution of the atoms around each residue fragment in the models and equivalent distributions derived from the database of known structures. In general, an average quality score below −5 indicates poorly packed structures. A score between −2.0 and −1.0 indicates protein threaded correctly. Profile3D judged the overall fold and side-chain packing of the models, while the general accuracy of the main chain dihedral angles has been checked with the Ramachandran plot using the program Procheck.

The final models chosen for the docking studies present scoring values close to those of the templates used and to the cytochromes 2C9 taken as reference structure for the X-ray solved mammalian p450s. In fact, the Ramachandran percentage of residues in the generally allowed regions is 94.7% for the models and 89–98% for the reference structures, the Profile 3D score is 170 for the models and 169–217 for the reference structures, and the Whatif score is −1.33 for the models and −0.66 to −1.36 for the reference structures.

The ligands have been built using the molecular editor of the program Quanta. After minimization they were docked into the active site, according to the best initial guess obtained by GOLD,[31] which corresponds approximately to the position of Warfarine into the 2C9.[27]

The tightness of the active site does not permit the compounds to make large rotational movements during the



pActivity = -10.17 IE/Surf -2.06
n = 10, $R^2 = 0.651$, $s^2 = 0.039$, $R^2_{cv} = 0.420$

pActivity = 12.36 $HB_{acceptor}$ − 1.68
n = 10, $R^2 = 0.771$, $s^2 = 0.026$, $R^2_{cv} = 0.619$

pActivity = -0.19 E_{HOMO} − 17.09
n = 10, $R^2 = 0.932$, $s^2 = 0.008$, $R^2_{cv} = 0.885$

Figure 6. Selected correlations obtained for the third series of ligands (CYP1A2 activity). Statistical parameters R = correlation coefficient, n = number of compounds, s = standard deviation are given. Assessment of the prediction power of the QSAR models is given by the cross-validated correlation coefficient R^2_{cv}.

Table 5. Interaction energies (kcal/mol) between the ligands of the second series and individual residues of *h*CYP1A2

Residues	Flavone	3-OH	5-OH	7-OH	3,7-OH	3,5,7-OH
R68	−6.08	−5.26	−4.81	−2.17	−0.01	−7.91
L71	−2.54	−0.18	−0.09	−1.30	−2.94	−0.16
T73	−1.41	−1.06	0	−0.63	−0.07	0
S74	−1.11	−6.82	0	−0.20	0	0
T84	−0.47	−0.62	−1.90	−1.19	−0.17	−0.66
F186	−4.84	−1.70	−2.22	−2.12	−0.81	−1.33
V187	−2.08	−0.28	−1.84	−1.70	−0.11	−2.65
N194	−0.02	−0.03	−1.62	−0.04	−2.24	−6.31
N272	−1.30	−0.85	−4.63	−0.14	−0.17	−0.28
D273	−2.28	−2.15	−1.81	−1.14	−0.90	−1.97
A277	−0.23	−0.21	−0.18	−0.14	−0.11	−0.32
D280	−2.18	−0.40	−1.46	−0.32	−0.05	−3.79
T345	−0.83	−0.40	−1.48	−1.54	−2.52	−6.96
I346	−2.85	−2.99	−3.91	−2.45	−4.97	−2.77
T458	−0.19	−6.45	−2.31	−8.46	−2.49	−5.45

Table 6. Interaction energies (kcal/mol) between the ligands of the second series and individual residues of *h*CYP1A1

Residues	Flavone	3-OH	5-OH	7-OH	3,7-OH	3,5,7-OH
R68	−0.94	−1.33	0.01	−0.18	−0.33	−0.11
L71	−0.63	−1.14	−1.66	−1.08	−1.57	−1.43
Y72	−2.15	−6.10	−2.82	−7.45	−5.57	−4.41
S84	−1.00	−0.96	−1.62	−9.62	−0.04	−1.70
F186	−3.01	−6.51	−3.88	−2.94	−3.27	−2.00
G187	−2.52	−2.22	−1.49	−0.12	−1.63	−0.22
V190	−0.61	−1.21	−0.09	−0.62	−1.56	−1.25
S192	−1.42	−8.36	0.00	−0.03	−1.57	0.00
D275	−1.65	−1.04	−7.24	−9.04	−7.05	−1.79
G278	−0.13	−0.50	−1.27	−0.58	−0.67	−1.60
A279	−1.24	−1.12	−2.46	−2.24	−1.38	−1.59
D282	−0.25	−9.63	−1.67	−3.42	−2.93	−5.65
T347	−2.15	−0.05	−2.06	−0.59	−7.23	−6.59
I348	−5.44	−1.56	−2.19	−0.58	−6.16	−7.32

dynamic simulations so we hypothesized that the side with which the inhibitors enter the active site, which might be determined by the position of the hydroxyl substituents, is fundamental for the inhibitory activity. To increase our confidence in this assumption we have performed various tests including a longer equilibration and/or simulation time and the use of explicit water molecules in a layer of 8 Å radius.

The molecular dynamics average coordinates of the complexes have been minimized and then the interaction energies of the inhibitor-complex have been evaluated.

MetaSite[17] has been used to derive a quantitative estimation of the hydrogen bond acceptor complementarity between the protein and the ligands. The progam is specially designed to predict the site of metabolism for xenobiotics starting from the 3D structure of a compound; however, the similarity indexes derived in order to rank the different atoms of the ligands can be exploited as descriptors for correlative purposes.

The geometries of the ligands studied were fully optimized by means of the ZINDO[18] program both in water solution and in chloroform to mimic the low dielectric environment of the enzyme-bounded ligands.

Acknowledgments

We thank the NFCR (National Foundation for Cancer Research) Centre for Drug Discovery, University of Oxford, UK, for financial support, and the FCT (Fundação para a Ciência e Tecnologia) for a doctoral scholarship for R.F.

References and notes

1. Hodek, P.; Trefil, P.; Stiborová, M. *Chem. Biol. Interact.* **2002**, *139*, 1.
2. Tsyrlov, B.; Mikhailenko, V. M.; Gelboin, H. V. *Biochim. Biophys. Acta-Protein Struct. Mol. Enzymol.* **1994**, *1205*, 325.
3. Breinholt, V. M.; Offord, E. A.; Brouwer, C.; Nielsen, S. E.; Brøsen, K.; Friedberg, T. *Food Chem. Toxicol.* **2002**, *40*, 609.
4. Rendic, S.; Di Carlo, F. J. *Drug Metab. Rev.* **1997**, *29*, 413.
5. Ferguson, L. R. *Mutat. Res.* **2001**, *475*, 89.
6. Lee, H.; Yeom, H.; Kim, Y. G.; Yoon, C. N.; Jin, C.; Choi, J. S.; Kim, B.-R.; Kim, D.-H. *Biochem. Pharmacol.* **1998**, *55*, 1339.
7. Moon, T.; Chi, M. H.; Kim, D.-H.; Yoon, C. N.; Choi, Y.-S. *Quant. Struct.-Act. Relat.* **2000**, *19*, 257.
8. Lewis, D. F. *Toxicology* **2000**, *144*, 197.
9. Lewis, D. F. *Inflammopharmacology* **2003**, *11*, 43.
10. De Rienzo, F.; Fanelli, F.; Menziani, M. C.; De Benedetti, P. G. *J. Comput. Aided Mol. Des.* **2000**, *14*, 93.
11. Szklarz, G. D.; Paulsen, M. D. *J. Biomol. Struct. Dyn.* **2002**, *20*, 155.
12. Da-Fonseca, R.; Menziani, M. C.; Melo, A.; Ramos, M. J. *Mol. Phys.* **2003**, *101*, 2731.
13. Lewis, D. F. V.; Lake, B. G.; Dickins, M.; Ueng, Y. F.; Goldfarb, P. S. *Xenobiotica* **2003**, *33*, 239.
14. Zhai, S.; Dai, R.; Wei, X.; Friedman, F. K.; Vestal, R. E. *Life Sci.* **1998**, *63*, PL119.
15. Zhai, S.; Dai, R.; Friedman, F. K.; Vestal, R. E. *Drug Metab. Dispos.* **1998**, *26*, 989.
16. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
17. Zamora, I.; Afzelius, L.; Cruciani, G. *J. Med. Chem.* **2003**, *46*, 2313.
18. Karelson, M. M.; Katritzky, A. R.; Zerner, M. C. *Int. J. Quant. Chem.* **1986**, *S20*, 521.
19. Menziani, M. C.; Montorsi, M.; De Benedetti, P. G.; Karelson, M. *Bioorg. Med. Chem.* **1999**, *7*, 2437.
20. Parikh, A.; Josephy, P. D.; Guengerich, F. P. *Biochemistry* **1999**, *38*, 5283.
21. Hadjokas, N. E.; Dai, R.; Friedman, F. K.; Spence, M. J.; Cusak, B. J.; Vestal, R. E.; Yongsheng, M. *Br. J. Pharmacol.* **2002**, *136*, 347.
22. Liu, J.; Ericksen, S. S.; Sivaneri, M.; Besspiata, D.; Fisher, C. W.; Szklarz, G. D. *Arch. Biochem. Biophys.* **2004**, *424*, 33.
23. Rost, B.; Sander, C. *Proteins* **1994**, *19*, 55.
24. Higgins, D.; Thompson, J.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673.
25. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
26. van Gunsteren, W. F.; Berendsen, J. C. *Mol. Phys.* **1977**, *34*, 1311.
27. Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Vinkovic, D. M.; Jhoti, H. *Nature* **2003**, *424*, 464.
28. Vriend, G. *J. Mol. Graphics* **1990**, *8*, 52.
29. Bowie, J. U.; Lüthy, R.; Eisenberg, D. A. *Science* **1991**, *253*, 164.
30. Laskowski, R. A.; Mac Arthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
31. Jones, G.; Willet, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.